Computer Physics Communications 125 (2000) 105–118

Computer Physics Communications

www.elsevier.nl/locate/cpc

Enhancements to the *GW* space-time method

L. Steinbeck [a,1], A. Rubio [b], L. Reining [c], M. Torrent [c,2], I.D. White [d], R.W. Godby [a]

[a] *Department of Physics, University of York, Heslington, York YO1 5DD, UK*
[b] *Departamento Fisica Teorica, Universidad de Valladolid, E-47011 Valladolid, Spain*
[c] *Laboratoire des Solides Irradiés, UMR 7642 CNRS - CEA/CEREM, Ecole Polytechnique, F-91128 Palaiseau, France*
[d] *Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge, CB3 0HE, UK*

Received 20 June 1999

Abstract

We describe the following new features which significantly enhance the power of the recently developed real-space imaginary-time *GW* scheme (Rieger et al., Comp. Phys. Commun. 117 (1999) 211) for the calculation of self-energies and related quantities of solids: (i) to fit the smoothly decaying time/energy tails of the dynamically screened Coulomb interaction and other quantities to model functions, treating only the remaining time/energy region close to zero numerically and performing the Fourier transformation from time to energy and vice versa by a combination of analytic integration of the tails and Gauss–Legendre quadrature of the remaining part and (ii) to accelerate the convergence of the band sum in the calculation of the Green's function by replacing higher unoccupied eigenstates by free electron states (plane waves). These improvements make the calculation of larger systems (surfaces, clusters, defects etc.) accessible.

PACS: 71.15.Th; 71.20.-b

Keywords: Electronic structure; Quasiparticle energies; Self-energy calculations; *GW* approximation

1. Introduction

Density-functional calculations provide reliable information about the ground state properties of electron systems but give, in principle, no access to the excitation spectrum of the system under study. Excitations can be described by many-body perturbation theory which is, however, at present only computationally feasible for real materials in its simplest form, the *GW* approximation of Hedin [1,2]. The latter gives a comparatively simple expression for the self-energy operator, which allows the one-particle Green's function of an interacting many-electron system to be described in terms of the Green's function of a hypothetical non-interacting system with an effective potential. The Green's function contains information not only about the ground-state density and energy but also about the quasiparticle (QP) spectrum. The *GW* approximation has been successfully applied to the calculation of QP bandstructures of semiconductors and other materials [3–6], for a recent review see Ref. [7].

[1] Present address: Institut für Festkörper- und Werkstofforschung, Postfach 270016, 01171 Dresden, Germany. E-mail: lutz@tmps16.mpg.tu-dresden.de.

[2] Present address: Commissariat à l'Energie Atomique-Centre d'Etudes de Bruyères-le-Chatel, BP12, 91680 Bruyères-le-Chatel, France.

0010-4655/00/$ – see front matter
PII: S0010-4655(99)00466-X

The real-space imaginary-time *GW* method, first proposed by Rojas et al. [8] and – in a revised form – described in detail by Rieger et al. [9] (we will refer to this paper as CPC I in the following) offers a more favourable scaling of the computational effort with system size than conventional reciprocal-space *GW* schemes [9]. It substantially reduces the computational effort and allows to study larger systems than previously possible without resorting to further approximations such as plasmon-pole models [3] for the energy dependence of the screened interaction or model dielectric functions [10].

The new features outlined in the present paper, particularly the new treatment of the (imaginary) time/energy dependence, further reduce the computational effort of the space-time *GW* scheme by almost an order of magnitude. This is achieved by fitting the smoothly decaying large energy/time tails of all quantities involved in a *GW* calculation to simple model functions and treating the remaining time/energy region numerically on a Gauss–Legendre grid rather than using an equidistant grid and fast Fourier transformations (FFT) from time to energy and vice versa. In the new scheme these Fourier transformations are performed by a combination of analytic integration of the tails and Gauss–Legendre quadrature of the remaining part. Another improvement of the method concerns the convergence of the calculated Green's function with the number of unoccupied eigenstates entering the eigenstate (band) sum in the Green's function Eq. (2.1) below. Higher unoccupied eigenstates are approximated by plane waves. This considerably reduces the number of eigenstates and energies which have to be computed in a density-functional calculation (usually within the local density approximation (LDA)) preceding a calculation of the self-energy with a given accuracy.

The present paper is organized as follows: first we give a brief summary of the real-space imaginary-time *GW* scheme in order to clarify notation in reference to CPC I (Section 2). Then we describe the new treatment of the time/energy dependence (Section 3) and the plane-wave substitution for accelerating the unoccupied-state sum convergence of the Green's function (Section 4).

2. Summary of the real-space imaginary-time *GW* method

In the real-space imaginary-time *GW* method [8,9] for computing electron self-energies and related quantities such as dielectric response functions and quasiparticle energies the basic quantities Green's function, dielectric response function, dynamically screened Coulomb interaction and self-energy are represented on a real-space grid and on the imaginary time axis. In those intermediate steps of the calculation where it is computationally more efficient to work in reciprocal space and imaginary energy we change to the latter representation by means of Fourier transforms. The choice of representing the time/energy dependence on the imaginary instead of on the real axis allows us to deal with smooth, decaying quantities which give faster convergence. To obtain the self-energy eventually on the real energy axis, we fit a model function to the computed self-energy on the imaginary axis, and continue it analytically to the real axis. The energy dependence of the dynamically screened interaction is fully taken into account within the method. The computational effort scales quadratically with the number of atoms in the unit cell and linearly with the number of energy points N_ω used to represent the energy dependence [9].

First, the zeroth-order Green's function is constructed in real space and imaginary time:

$$G_{\mathrm{LDA}}(\mathbf{r},\mathbf{r}';i\tau)=\begin{cases} i\sum_{n\mathbf{k}}^{\mathrm{occ}}\Psi_{n\mathbf{k}}(\mathbf{r})\Psi_{n\mathbf{k}}^{*}(\mathbf{r}')\exp(\varepsilon_{n\mathbf{k}}\tau), & \tau>0,\\ -i\sum_{n\mathbf{k}}^{\mathrm{unocc}}\Psi_{n\mathbf{k}}(\mathbf{r})\Psi_{n\mathbf{k}}^{*}(\mathbf{r}')\exp(\varepsilon_{n\mathbf{k}}\tau), & \tau<0,\end{cases} \tag{2.1}$$

from the LDA wavefunctions $\Psi_{n\mathbf{k}}(\mathbf{r})$ and eigenvalues $\varepsilon_{n\mathbf{k}}$. Then the RPA irreducible polarizability is formed in real space and imaginary time:

$$\chi^{0}(\mathbf{r},\mathbf{r}';i\tau)=-iG_{\mathrm{LDA}}(\mathbf{r},\mathbf{r}';i\tau)G_{\mathrm{LDA}}(\mathbf{r}',\mathbf{r};-i\tau), \tag{2.2}$$

and Fourier transformed to reciprocal space and imaginary energy and the symmetrised dielectric matrix [11] is constructed in reciprocal space,

$$\tilde{\varepsilon}_{\mathbf{G}\mathbf{G}'}(\mathbf{k}, i\omega) = \delta_{\mathbf{G}\mathbf{G}'} - \frac{4\pi}{|\mathbf{k}+\mathbf{G}|\,|\mathbf{k}+\mathbf{G}'|}\,\chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{k}, i\omega). \tag{2.3}$$

After that the symmetrised dielectric matrix is inverted for each $\mathbf{k}$ point and each imaginary energy in reciprocal space and the screened Coulomb interaction is calculated:

$$W_{\mathbf{G}\mathbf{G}'}(\mathbf{k}, i\omega) = \frac{4\pi}{|\mathbf{k}+\mathbf{G}|\,|\mathbf{k}+\mathbf{G}'|}\,\tilde{\varepsilon}^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{k}, i\omega), \tag{2.4}$$

and Fourier transformed to real space and imaginary time. From that the self-energy operator

$$\Sigma(\mathbf{r}, \mathbf{r}'; i\tau) = iG_{\mathrm{LDA}}(\mathbf{r}, \mathbf{r}'; i\tau)W(\mathbf{r}, \mathbf{r}'; i\tau), \tag{2.5}$$

and its expectation values $\langle \mathbf{q}n|\Sigma(i\tau)|\mathbf{q}n\rangle$ are computed. The latter are Fourier transformed to imaginary energy and fitted to a model function allowing analytic continuation onto the real energy axis and evaluation of the quasiparticle corrections to the LDA eigenvalues by first-order perturbation theory in $\langle \Sigma - V_{\mathrm{xc}}^{\mathrm{LDA}}\rangle$. Since all quantities go to zero with increasing $|\mathbf{r}-\mathbf{r}'|$ we use a finite cutoff region in real space which we call the interaction cell. Further details of the method can be found in CPC I.

3. New treatment of time/energy dependence

3.1. Motivation and basic idea

The functions we are dealing with are relatively smooth on the imaginary time/energy axis. This allows to employ a regular time/energy grid which has the advantage that the Fourier transformation from imaginary time to imaginary energy and vice versa can be done efficiently by fast Fourier transformation (FFT). However, we still need of the order of 100 grid points [3] for good convergence (resulting quasiparticle energies converged within 30 meV with respect to the τ/ω grid parameters). This point is illustrated by Fig. 1 showing the matrix element of the correlation self-energy [4] for the uppermost valence band of Si at Γ, calculated (with $\Delta\omega = 0.17$ Hartree) with 30, 60, and 120 FFT grid points (a) on the imaginary energy axis and (b) analytically continued to real energies. Crucially, the convergence on the imaginary axis transforms into a convergence of similar quality on the real axis upon analytic continuation. Looking at the time/energy behaviour of the key quantities, particularly those which have to be Fourier transformed such as polarizability, screened interaction and the matrix elements of the self-energy (see Fig. 2) we observe that they possess nontrivial structure only in the region close to $i\tau = 0$ ($i\omega = 0$) whereas they decay smoothly to zero for large imaginary times or energies. The FFT grid has to be large enough to take account of the tails (reduce aliasing) and at the same time it needs to be sufficiently dense to describe the structure in the region close to the origin properly.

This suggests another approach: represent the functions on a suitably chosen grid in a fixed and comparatively small time/energy interval and fit the large $i\tau/i\omega$ tails to simple model functions which can be Fourier transformed analytically, a method suggested by Blase et al. [12] in the context of their earlier mixed-space method [13]. For the part handled numerically we choose a Gauss–Legendre (GL) grid (linearly transformed from $(-1, 1)$ to $(0, \tau_{\max})$ or $(0, \omega_{\max})$, respectively). This turns out to be very efficient since the functions have to be computed for a relatively small number of time or energy points only and this computation of the functions is much more time-consuming than the Fourier transformations themselves which are done by Gaussian quadrature over the numerical values and

[3] If we were working on the real energy axis this number would be an order of magnitude larger still, as can be concluded from Ref. [16].

[4] Defined as the difference between the *GW* self-energy and its time-independent bare-exchange part.

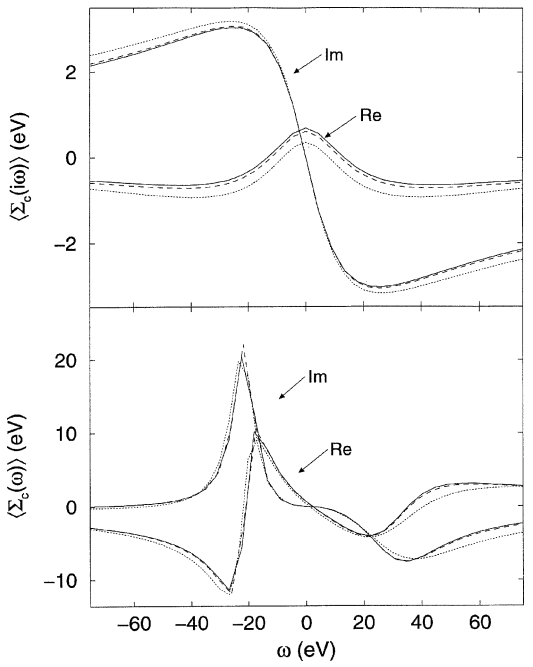

Fig. 1.

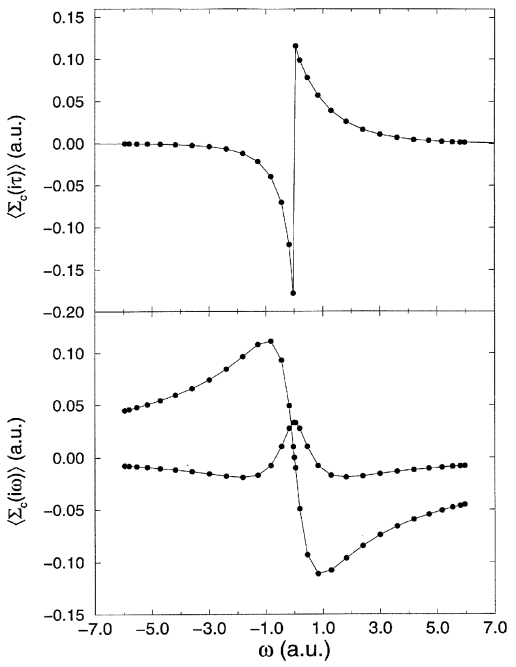

Fig. 2.

Fig. 1. Convergence of the matrix element of the correlation self-energy $\langle \Sigma_c(i\omega)\rangle$ for the valence-band top $\Gamma_{25}^{v\prime}$ of silicon with respect to ω_{max} with a fixed energy grid spacing of $\Delta\omega = 0.17$ Har. The top panel shows the calculated self-energy on the imaginary axis with the analytically continued dependence on real energy shown in the lower panel. The lines correspond to $\omega_{max} = 5$ Har (dotted), 10 Har (dashed) and 20 Har (solid), i.e. 30, 60, and 120 FFT grid points, respectively.

Fig. 2. Matrix elements $\langle \Sigma_c(i\tau)\rangle$ (top panel) and $\langle \Sigma_c(i\omega)\rangle$ (lower panel) of the correlation self-energy for the valence-band top $\Gamma_{25}^{v\prime}$ of silicon. The Gauss–Legendre grid points for $\tau_{max} = \omega_{max} = 6$ a.u. and $N_\tau = N_\omega = 15$ are indicated by filled circles. This figure gives an idea about a typical functional form on the imaginary time (energy) axis and a typical Gauss–Legendre grid which has to be used to Fourier transform this function with good convergence. The other quantities which have to be Fourier transformed from time to energy or vice versa have similar (or less) structure.

adding the Fourier transform of the tail [5]. The fit of the tail only needs to be performed whenever a quantity has to be Fourier transformed from $i\tau$ to $i\omega$ or vice versa. The rest of the calculation is restricted to the GL grid. Hence the following quantities have to be fitted: (1) the polarizability $\chi^0_{\mathbf{GG}'}(\mathbf{k}, i\tau)$, (2) the screened Coulomb interaction $W_{\mathbf{GG}'}(\mathbf{k}, i\omega)$ and (3) the matrix elements of the correlation part of the self-energy $\langle \mathbf{q}n|\Sigma_c(i\tau)|\mathbf{q}n\rangle$.

The advantages of this treatment of the time/energy dependence are obvious: all quantities have to be computed for a much smaller number of imaginary times/energies thus saving computational time and reducing storage requirements while retaining the flexibility to accomodate general functional forms of the energy dependence and not being restricted to particular forms such as plasmon-pole models.

3.2. Polarizability

The imaginary-time dependence of the Green's function derives from the the decaying exponentials $\exp(\varepsilon_{n\mathbf{k}}\tau)$ with $\varepsilon_{n\mathbf{k}} < 0$ (> 0) for $\tau > 0$ (< 0) in Eq. (2.1). The asymptotic behavior at large imaginary times is determined

[5] An alternative to a GL grid would be a grid which is more dense towards the origin by a suitable transformation of variables. However, such grids do not perform well in computing the 'fast-oscillating' components, i.e. $i\int d\tau f(i\tau)\exp(-i\omega\tau)$ for large ω.

by the slowest-decaying exponentials and can thus be approximated by a single exponential. This asymptotic imaginary-time dependence carries over to the polarizability Eq. (2.2). For that reason we fit the large-$i\tau$ tails of each ($\mathbf{G}\mathbf{G}'\mathbf{k}$) element of the polarizability $\chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{k}, i\tau)$ [6] to a decaying exponential $a\exp(-b\tau)$ (with $b > 0$ and for $\tau \geqslant 0$ only since χ^0 is symmetric in τ). The two fit parameters a and b are exactly determined by fitting two time points: the outermost GL grid point and one additional point at $1.3\tau_{\max}$. This fitting procedure turns out to be very reliable.

The Fourier transformation from $i\tau$ to $i\omega$ is done in the following way:

$$\begin{aligned}\chi^0(i\omega_j) &= \sum_{i=-i_{\max}}^{i_{\max}} p_i\big[\bar{\chi}^0(i\tau_i) - a\exp\big(-b|\tau_i|\big)\big]\exp(-\omega_j\tau_i) + \int_{-\infty}^{\infty} d\tau\, a\exp\big(-b|\tau|\big)\exp(-\omega_j\tau)\\ &= 2\sum_{i=1}^{i_{\max}} p_i\big[\bar{\chi}^0(i\tau_i) - a\exp\big(-b|\tau_i|\big)\big]\cos(\omega_j\tau_i) + \frac{2a}{b^2+\omega_j^2}\end{aligned} \tag{3.1}$$

with GL grid points τ_i and ω_j, GL weights p_i, fit parameters a and b, and $\bar{\chi}^0(i\tau) = -i\chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{k}, i\tau)$.

For a small number of matrix elements $\chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{k}, i\tau)$ (typically less than 5% of all matrix elements as long as $\tau_{\max}$ is large enough to accomodate all the nontrivial structure of $\chi_0(i\tau)$) the large $i\tau$ tails cannot be fitted to a decaying exponential because they increase or change sign. This is only the case for small matrix elements where the function is already close to zero at $\tau_{\max}$ anyway. We set $\chi^0(1.3\tau_{\max})$ to $0.1\chi_0(\tau_{\max})$ there, i.e. choose a reasonable decaying constant which takes the (already small) function smoothly to zero. Simply setting the matrix element to zero for $\tau > \tau_{\max}$ would render the ensuing fit of $W_{\mathbf{G}\mathbf{G}'}(\mathbf{k}, i\omega)$ unneccessarily difficult. Anyhow, $\tau_{\max}$ is a convergence parameter which can be varied to check the quality of the results.

3.3. Dynamically screened Coulomb interaction

The large-imaginary-energy tail of the dynamically screened interaction $W_{\mathbf{G}\mathbf{G}'}(\mathbf{k}, i\omega)$ is fitted to the Fourier transform of a decaying exponential

$$\int_{-\infty}^{\infty} d\tau\, a\exp\big(-b|\tau|\big)\exp(\pm i\omega\tau) = \frac{2ab}{b^2+\omega^2} = \frac{\alpha}{\beta^2+\omega^2}. \tag{3.2}$$

The energy region where W is treated numerically has to be large enough to comprise the nontrivial structure of $W(i\omega)$. We found that $\omega_{\max}$ should be between 3 and 10 times the plasmon energy [7] for good convergence. We could perform the tail fit along similar lines as that of χ_0, i.e. subtract the analytic tail function from the given imaginary-energy W in $(0, \omega_{\max})$, Fourier transform this difference numerically and add the analytically given Fourier transform of the function fitted to the tail back in. However, for a large number of matrix elements $W_{\mathbf{G}\mathbf{G}'}(\mathbf{k}, i\omega)$ the tail fit yields a negative β^2 because they decay more rapidly than $1/\omega^2$. In this case the function $\alpha/(\beta^2+\omega^2)$ has a pole inside the interval $(0, \omega_{\max})$ which does not allow the analytic Fourier transformation to be performed and which makes the numerical Fourier transformation of the difference between W and the fit function virtually impossible to compute. That is why we integrate the analytic tail function from $\omega_{\max}$ to ∞ in this case, provided that $\beta^2 > -\omega^2_{\max}$. Part of this integral can still be solved analytically whereas the remainder is treated numerically. The Fourier transform $\overline{W}(i\tau) = -iW_{\mathbf{G}\mathbf{G}'}(\mathbf{k}, i\tau)$ is given by:

[6] In our experience fitting in reciprocal space rather than in real space is both easier and more efficient. For very large systems where it is advantageous to use a real-space representation throughout [9] the fit can of course also be done in real space.

[7] Taking the free-electron-gas plasmon energy at the average valence density of bulk Si that corresponds to between 2 and 6 Hartree.

$$\overline{W}(i\tau_i) = \frac{1}{\pi}\sum_{j=1}^{j_{\max}} p_j\, W(i\omega_j)\cos(\omega_j\tau_i) + \frac{1}{\pi}\int_{\omega_{\max}}^{\infty} d\omega\, \frac{\alpha}{\beta^2+\omega^2}\cos(\omega\tau_i) \tag{3.3}$$

with

$$\frac{1}{\pi}\int_{\omega_{\max}}^{\infty} d\omega\, \frac{\alpha}{\beta^2+\omega^2}\cos(\omega\tau) = -\frac{\alpha}{\pi}\int_{\omega_{\max}}^{\infty} d\omega\, \frac{\beta^2}{\beta^2+\omega^2}\,\frac{\cos(\omega\tau)}{\omega^2} + \frac{\alpha}{\pi}\left[\frac{\cos(\omega_{\max}\tau)}{\omega_{\max}} - \tau\,\mathrm{Si}(\omega_{\max}\tau)\right] \tag{3.4}$$

with the sine integral

$$\mathrm{Si}(\omega_{\max}\tau) = \int_{\omega_{\max}}^{\infty} d\omega\, \frac{\sin(\omega\tau)}{\tau}. \tag{3.5}$$

Here τ_i and ω_j are the GL grid points, p_j the GL weights, α and β^2 the fit parameters and $\omega_{j_{\max}}$ the outermost GL grid point. The integral on the right hand side of Eq. (3.4) is solved numerically using a transformed GL grid. It converges rapidly since the integrand is going to zero like $1/\omega^4$. The second part of Eq. (3.4) is given analytically. In this way most $W_{\mathbf{GG}'}(\mathbf{k}, i\omega)$ can be fitted except for a small number where $\beta^2 \leqslant -\omega_{\max}^2$ (this only occurs for matrix elements which are small anyway). In the latter case we take the correct value at $\omega_{\max}$ smoothly to zero by setting β^2 to $-0.9\omega_{\max}^2$. Again, the quality of the results can be checked by varying $\omega_{\max}$.

3.4. Matrix elements of correlation self-energy

The matrix elements of the correlation self-energy are fitted in a similar way as the polarizability. As the asymptotic time dependence of the self-energy is determined by that of the Green's function which is 'shorter-ranged' (in imaginary time) than W we again fit to a decaying exponential $a\exp(-b\tau)$. This time, however, we have to perform separate fits on the positive and negative half-axis since the self-energy is not symmetric in imaginary time. Therefore we obtain two contributions to $\langle\Sigma_{\mathrm{c}}(i\omega)\rangle$ for positive imaginary energies (with $\overline{\Sigma}_{\mathrm{c}}(i\tau) = -i\langle\Sigma_{\mathrm{c}}(i\tau)\rangle$):

$$\begin{aligned}\big\langle\Sigma_{\mathrm{c}}(i\omega_j)\big\rangle = {} & \sum_{i=1}^{i_{\max}} p_i\big[\overline{\Sigma}_{\mathrm{c}}(i\tau_i) - a_+\exp\big(-b_+|\tau_i|\big)\big]\exp(-i\omega_j\tau_i) + \frac{a_+}{b_+ + i\omega_j} \\ & + \sum_{i=-1}^{-i_{\max}} p_i\big[\overline{\Sigma}_{\mathrm{c}}(i\tau_i) - a_-\exp\big(-b_-|\tau_i|\big)\big]\exp(-i\omega_j\tau_i) + \frac{a_-}{b_- + i\omega_j},\end{aligned} \tag{3.6}$$

here τ_i and ω_j denote GL grid points, p_i GL weigths, and a_+, a_-, b_+, and b_- are the fit parameters. The self-energy matrix elements on the negative imaginary-energy half-axis are given by $\langle\Sigma_{\mathrm{c}}(-i\omega)\rangle = \langle\Sigma_{\mathrm{c}}(i\omega)\rangle^*$.

All matrix elements of Σ_{c} could be fitted in this way for all the systems investigated so far (Si, Ge, GaN).

3.5. Tests for bulk silicon and zincblende GaN

In order to test the stability and accuracy of the tail fit and the convergence with the GL grid size we performed self-energy calculations for bulk Si (cutoff parameters are given in Table 1 and further details are as in CPC I). Fig. 2 gives an idea of the sampling of the matrix element of the correlation self-energy for the uppermost valence band at Γ on the imaginary time and imaginary energy axis by a GL grid with $\tau_{\max} = \omega_{\max} = 6$ a.u. and $N_\tau = N_\omega = 15$. The convergence of two typical quasiparticle energies with the number of GL grid points and the size of the imag-

inary time/energy region treated explicitly is summarized in Table 2 [8]. We chose the lowest conduction state $\Gamma_{15}^{c\prime}$ and the valence-band bottom Γ_1^v. The latter is in our experience the slowest-converging quasiparticle energy which is most sensitive to details of the calculation and cutoff parameters. Obviously, $N_\tau = N_\omega$ and $\tau_{max} = \omega_{max}$ have to be increased simultaneously since the main structure of the functions we are dealing with is restricted to a limited region around the origin. We observe that 15 GL grid points and $\tau_{max} = \omega_{max} = 5$ a.u. are sufficient to converge even the slowest-converging valence-band bottom to within 30 meV. The same accuracy is obtained with FFT grids for $\tau_{max} = \omega_{max} = 20$ a.u. and 120 grid points, resulting in QP energies of 3.22 eV ($\Gamma_{15}^{c\prime}$) and −11.58 eV (Γ_1^v).

Table 1
Cutoff and grid parameters used for the test calculations for Si and zincblende GaN, respectively, in Section 3 of the present work

Parameter	Si	GaN
LDA plane-wave cutoff (in Ry)	13.5	50.
GW plane-wave cutoff[a] (in Ry)	19.	38.
GW real-space grid	$9 \times 9 \times 9$	$12 \times 12 \times 12$
Band cutoff (in Ry)	10.	16.
Size of **k** grid	$4 \times 4 \times 4$	$4 \times 4 \times 4$
Range τ_{max} of time grid[b] (in a.u.)	6.	6.
Size of time (energy) grid[b]	15	15

[a] Energy cutoff corresponding to radius of circumscribing sphere, see Ref. [9].
[b] This parameter is varied in the tests of Section 3.

Table 2
Convergence of quasiparticle energies $\Gamma_{15}^{c\prime}$ and Γ_1^v (in eV, top of valence band has been set to zero) for Si with respect to Gauss–Legendre grid region $\tau_{max} = \omega_{max}$ (in a.u.) and number of grid points $N_\tau = N_\omega$

$\tau_{max} = \omega_{max}$	$N_\tau = N_\omega$					
	10	12	15	18	21	25
$\Gamma_{15}^{c\prime}$						
3.	3.27	3.28	3.28			
4.	3.24	3.24	3.24			
5.		3.23	3.23	3.23		
6.			3.23	3.23	3.22	
7.				3.22	3.23	3.22
Γ_1^v						
3.	−11.54	−11.53	−11.52			
4.	−11.58	−11.52	−11.52			
5.		−11.53	−11.56	−11.56		
6.			−11.56	−11.58	−11.60	
7.				−11.63	−11.58	−11.59

[8] These quasiparticle energies are slightly different from those given in CPC I since head and wings of the dielectric matrix have not been computed separately using a different **k** grid in the present paper.

Table 3
Convergence of quasiparticle energies (in eV, top of valence band has been set to zero) at the Γ and X point of zincblende GaN with respect to Gauss–Legendre grid region $\tau_{\max} = \omega_{\max}$ (in a.u.) and number of time/energy grid points $N_\tau = N_\omega$

$N_\tau = N_\omega$	12	15	18	25
$\tau_{\max} = \omega_{\max}$	4.	5.	6.	8.
Γ_1^{v}	−15.88	−15.95	−15.93	−15.94
Γ_1^{c}	3.03	3.00	3.00	2.99
Γ_{15}^{c}	11.56	11.54	11.53	11.52
X_1^{v}	−12.96	−13.00	−12.97	−12.98
X_3^{v}	−6.38	−6.39	−6.38	−6.38
X_5^{v}	−2.66	−2.66	−2.66	−2.66
X_1^{c}	4.43	4.41	4.40	4.39
$X_1^{\mathrm{c}'}$	7.91	7.88	7.87	7.86
X_3^{c}	13.23	13.19	13.16	13.14
X_5^{c}	15.32	15.28	15.26	15.24

Our *GW* calculations for GaN in the zincblende structure were carried out at the experimental lattice constant (a = 8.54 a.u.). The LDA wavefunctions and eigenvalues used in the self-energy calculation were obtained from a standard plane-wave pseudopotential calculation. The Ga 4s and 4p states and the *N* 2s and 2p states were treated as valence states and the soft pseudopotentials of Troullier and Martins [14] were used. The cutoff parameters are given in Table 1. Table 3 shows the convergence of the resulting QP energies at Γ and X as a function of the time/energy GL grid parameters. From these results we conclude that the speed of convergence of the QP energies of GaN with respect to the imaginary time/energy grid is similar to that found for bulk Si.

The agreement of our calculated quasiparticle energies for Si and GaN with experiment is similar to that of previous *GW* calculations but – in contrast to these earlier works – dynamical effects are fully included here.

Fitting of the tails allows to reduce the time/energy region which is treated explicitly by more than a factor of three. This saves the same factor in the number of FFT grid points. Employing a GL grid enables us to reduce the number of points where the functions have to be computed by another factor between two and three. In total, the CPU time, memory and disk space requirements decrease by a factor of seven to eight in comparison with the time/energy FFT grid treatment described in CPC I.

4. Plane wave substitution

4.1. Motivation and basic idea

A large number of unoccupied states have to be included in the band sum in the Green's function Eq. (2.1) for a proper convergence of the resulting self-energy and QP energies. With growing system size it becomes increasingly difficult to provide such a large number of eigenstates by a density-functional calculation since direct diagonalization may not be computationally feasible whereas iterative diagonalization yields only a limited number of eigenstates or becomes prohibitively expensive. Besides that it would be desirable to accelerate the convergence of the unoccupied-state sum in order to reduce the computational effort for the calculation of the Green's function.

On the other hand we expect that the higher the energy of an unoccupied state the better it should be approximated by a free-electron state (plane wave). This is illustrated by Fig. 3 showing the band energy as a function of the band number for (LDA) eigenstates of Si and the corresponding plane-wave states with wavevectors $\mathbf{K} = \mathbf{k} + \mathbf{G}$, $\mathbf{G}$ being reciprocal lattice vectors of Si. At higher energies the two spectra look remarkably similar if we allow for a constant

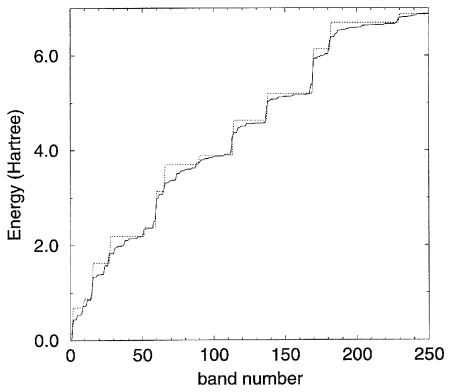

Fig. 3. Band energies as a function of band number for the LDA eigenstates of bulk Si (solid line) and the corresponding plane-wave states (dotted line). The latter have been shifted upwards by 0.26 Ry (with this shift the energies of LDA and PW state number 50 coincide). At higher energies the two spectra look remarkably similar (see text).

energy shift between the two. Although closer examination shows that this assumption is not fully justified for states with moderately high energies, which, for symmetry reasons, rather resemble linear combinations of several plane waves, the sum of all unoccupied states above a certain energy cutoff can still be reasonably well described by a corresponding sum of plane waves. The aspect of taking proper account of the weight of the higher unoccupied states seems to be more important than their explicit form. Thus there is good reason to expect that the number of unoccupied eigenstates which have to be explicitly included into the band sum in Eq. (2.1) can be substantially reduced by adding a sum of plane waves replacing the omitted higher unoccupied states [9]. This is indeed the case as is demonstrated by the significant improvement of the convergence of the QP energies as a function of the band cutoff upon adding a plane-wave (PW) contribution to the Green's function, cf. Figs. 4 and 5.

4.2. Method

The PW contribution to the Green's function Eq. (2.1) takes the following form in real-space:

$$\Delta G_{\mathrm{PW}}(\mathbf{r},\mathbf{r}';i\tau)=-\frac{i}{V}\sum_{\mathbf{K}}\exp(i\mathbf{K}\mathbf{r})\exp(-i\mathbf{K}\mathbf{r}')\exp\left(-\frac{\tau}{2}\left(\mathbf{K}^2-k_0^2\right)\right), \tag{4.1}$$

where the $\mathbf{K}$ vectors corresponding to energies below the lowest PW energy are excluded from the reciprocal-space sum. The plane waves are normalized with respect to the crystal volume $V = V_{\mathrm{UC}}N_{\mathbf{k}}$, with $N_{\mathbf{k}}$ being the number of $\mathbf{k}$ points in the Brillouin zone (BZ) and V_{UC} the volume of the unit cell. The energies of the PW states are measured with respect to an energy zero $\Delta E = k_0^2/2$ which is determined by adjusting the energy of the highest LDA eigenstate included and the highest PW state not included in the calculation of the Green's function, see Eq. (4.4) below. ΔG_{PW} can be computed analytically by transforming the $\mathbf{K}$ sum into an integral by $\sum_{\mathbf{K}} \to V_{\mathrm{UC}}/(2\pi)^3 \int d^3K$, i.e. taking the limit $N_{\mathbf{K}} \to \infty$ and solving the resulting integral. It turns out, however, that it is more practical to compute ΔG_{PW} numerically instead, even though that is computationally slightly more expensive. In this way the contributions of the (LDA) eigenstates of the system and the plane waves are treated on an equal footing which makes for a smoother convergence because of compensation of errors arising from discretization. Fourier transformation of $\Delta G_{\mathrm{PW}}(\mathbf{r},\mathbf{r}';i\tau)$ to reciprocal space results in:

$$\Delta G_{\mathrm{PW}}(\mathbf{k},\mathbf{G},\mathbf{G}';i\tau)=\frac{1}{N_{\mathbf{k}}V_{\mathrm{UC}}}\exp\left(-\frac{\tau}{2}k_0^2\right)\exp\left(-\frac{\tau}{2}(\mathbf{k}+\mathbf{G})^2\right)\delta_{\mathbf{G}\mathbf{G}'}, \tag{4.2}$$

[9] A related approach was followed by James and Woodley [15].

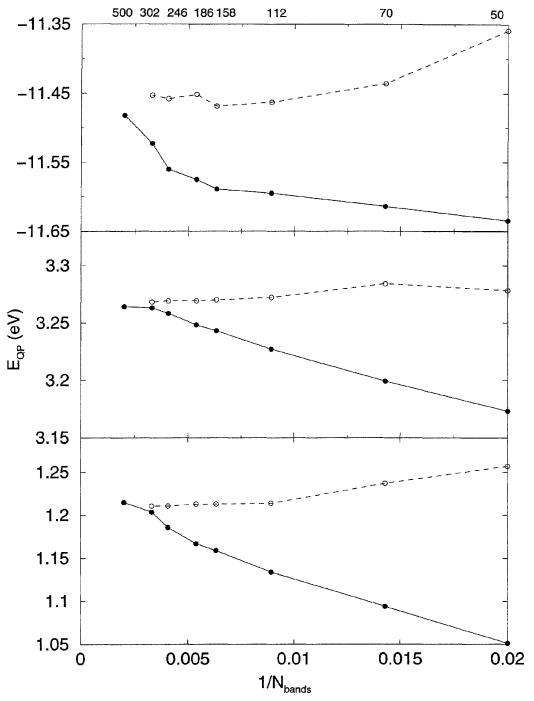

Fig. 4.

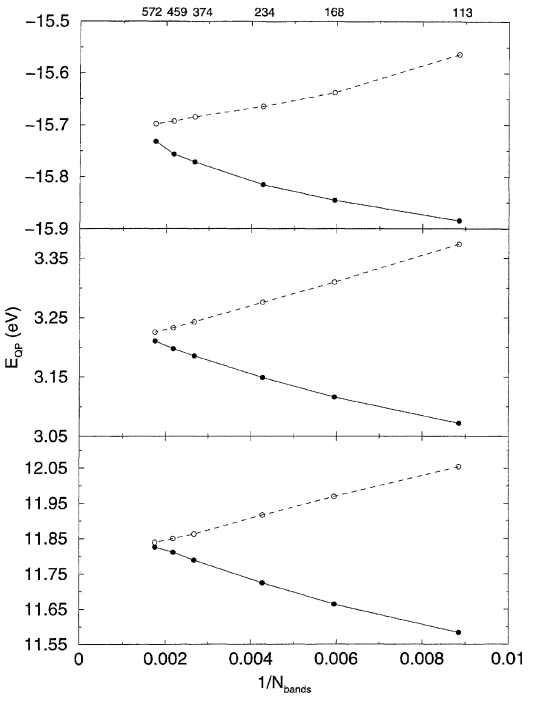

Fig. 5.

Fig. 4. Calculated valence band bottom Γ_1^{v} (top panel, the valence band maximum has been set to zero), direct gap $\Gamma_{15}^{\mathrm{c}\prime}$ (center), and minimal gap of bulk Si as a function of the inverse of N_{bands}, the number of LDA eigenstates used for the calculation of the Green's function. Two sets of data are shown, the filled circles (solid lines) refer to calculations with LDA eigenstates only whereas the open circles (dashed lines) are the results of corresponding calculations where the PW contribution was added (see text). The numbers along the top axis are the N_{bands} values used in the respective calculations. The number of LDA eigenstates needed to converge the quasiparticle energy is significantly reduced by adding the PW contribution.

Fig. 5. Same as Fig. 4 for valence band bottom Γ_1^{v} (top panel), direct gap Γ_1^{c} (center), and conduction state Γ_{15}^{c} (bottom) of zincblende GaN.

with reciprocal-lattice vectors $\mathbf{G},\mathbf{G}'$ of the system with $|\mathbf{k}+\mathbf{G}|^2/2$ larger than a given cutoff energy. As Eq. (4.2) is diagonal in $\mathbf{G}$ it is more efficient to first set up ΔG_{PW} in reciprocal space and then transform it to real space before adding it to the Green's function contribution of the LDA eigenstates taken into account explicitly.

In order not to destroy the crystal symmetry of the Green's function when adding the PW contribution only complete stars of $\mathbf{G}$ vectors (groups of plane-wave states energetically degenerate at any $\mathbf{k}$ in the BZ) must be included into or excluded from ΔG_{PW}. On the other hand the number of $\mathbf{G}$ excluded from ΔG_{PW} should be equal to N_{bands}, the number of LDA eigenstates taken into account explicitly, which, in turn, has to be determined in such a way as to include complete groups of energetically degenerate (at any $\mathbf{k}$) LDA eigenstates. These two demands cannot in general be fulfilled simultaneously. As a compromise we calculate ΔG_{PW} from PW states $(\mathbf{k}+\mathbf{G}_i)$ (ordered with respect to their energy) with weights

$$W_i(\mathbf{k}) = \begin{cases} 0 & i \leqslant N_1(\mathbf{k}), \\ 1 - \dfrac{N_{\mathrm{bands}} - N_1(\mathbf{k})}{N_2(\mathbf{k}) - N_1(\mathbf{k})} & N_1(\mathbf{k}) < i \leqslant N_2(\mathbf{k}), \\ 1 & i > N_2(\mathbf{k}), \end{cases} \tag{4.3}$$

Table 4
Cutoff and grid parameters used for the test calculations for Si and zincblende GaN, respectively, in Section 4 of the present work

Parameter	Si	GaN
LDA plane-wave cutoff (in Ry)	19.	50.
GW plane-wave cutoff[a] (in Ry)	26.	50.
GW real-space grid	$12 \times 12 \times 12$	$15 \times 15 \times 15$
Band cutoff[b] (in Ry)	10.	27.
Size of **k** grid	$4 \times 4 \times 4$	$4 \times 4 \times 4$
Range τ_{max} of time grid (in a.u.)	6.	6.
Size of time (energy) grid	15	15

[a] Energy cutoff corresponding to radius of circumscribing sphere, see Ref. [9].
[b] This parameter is varied in the tests of Section 4.

Table 5
Calculated quasiparticle energies at the Γ and X point for Si (in eV) as a function of the number of LDA eigenstates N_{bands} included in the calculation of the Green's function. Two sets of results are shown, obtained with including (b) or not including (a) the PW contribution to the Green's function (see text). The valence band maximum has been set to zero

Band cutoff [Ry]		6.	8.	12.	14.	16.	18.
N_{bands}		70	112	186	246	302	500
Γ_1^v	(a)	−11.61	−11.60	−11.58	−11.56	−11.52	−11.48
	(b)	−11.44	−11.46	−11.45	−11.46	−11.45	
$\Gamma_{15}^{c\prime}$	(a)	3.20	3.23	3.25	3.26	3.26	3.26
	(b)	3.28	3.27	3.27	3.27	3.27	
$\Gamma_2^{c\prime}$	(a)	3.93	3.94	3.95	3.96	3.96	3.96
	(b)	3.92	3.94	3.95	3.96	3.96	
X_1^v	(a)	−7.71	−7.69	−7.68	−7.66	−7.64	−7.63
	(b)	−7.64	−7.65	−7.64	−7.64	−7.64	
X_4^v	(a)	−2.85	−2.83	−2.82	−2.81	−2.80	−2.79
	(b)	−2.78	−2.80	−2.79	−2.80	−2.79	
X_1^c	(a)	1.24	1.27	1.31	1.33	1.35	1.36
	(b)	1.38	1.36	1.36	1.35	1.35	
X_4^c	(a)	10.75	10.75	10.75	10.74	10.72	10.71
	(b)	10.67	10.70	10.70	10.70	10.70	

where $N_1(\mathbf{k})$ and $N_2(\mathbf{k})$ are the largest possible/smallest possible total number of PW states smaller than/larger than N_{bands}, respectively, both containing complete stars of PW states only. This **k** dependent cutoff ensures preservation of both symmetry and number of bands and works well in practice.

In order to account for the difference in the energy spectra of LDA eigenstates and plane waves we introduce an energy shift

$$\Delta E = \frac{1}{2}\left|G(N_{\text{bands}})\right|^2 - E_{\text{cut}} + E_{\text{VBB}} \tag{4.4}$$

between the highest PW state(s) not included and the highest LDA eigenstate(s) included in the band sum in Eq. (2.1). E_{cut} and E_{VBB} stand for the energy (at $\mathbf{k} = 0$) of the highest LDA state taken into account and the LDA valence-band bottom, respectively (both measured with respect to the Fermi level which is chosen halfway between valence band top and conduction band bottom) and $\frac{1}{2}|G(N_{\text{bands}})|^2$ is the energy of plane wave number N_{bands} at $\mathbf{k} = 0$. Taking this energy shift at $\mathbf{k} = 0$ is somewhat arbitrary, but it turns out that the resulting self-energies and QP energies are not sensitive to the exact value of the energy shift [10].

Table 6
Same as Table 5 for zincblende GaN

Band cutoff [Ry]		12.	16.	20.	28.	32.	36.
N_{bands}		113	168	234	374	459	572
Γ_1^{v}	(a)	−15.88	−15.85	−15.82	−15.77	−15.76	−15.73
	(b)	−15.56	−15.64	−15.66	−15.68	−15.69	−15.70
Γ_1^{c}	(a)	3.07	3.12	3.15	3.19	3.20	3.21
	(b)	3.38	3.31	3.28	3.24	3.23	3.23
Γ_{15}^{c}	(a)	11.58	11.66	11.72	11.79	11.81	11.83
	(b)	12.05	11.97	11.92	11.86	11.85	11.84
X_1^{v}	(a)	−12.93	−12.91	−12.90	−12.87	−12.86	−12.85
	(b)	−12.75	−12.81	−12.82	−12.83	−12.84	−12.84
X_3^{v}	(a)	−6.39	−6.34	−6.30	−6.26	−6.24	−6.22
	(b)	−6.10	−6.15	−6.17	−6.20	−6.20	−6.21
X_5^{v}	(a)	−2.66	−2.63	−2.61	−2.59	−2.58	−2.57
	(b)	−2.52	−2.54	−2.55	−2.56	−2.57	−2.57
X_1^{c}	(a)	4.46	4.55	4.61	4.67	4.70	4.72
	(b)	4.94	4.86	4.81	4.76	4.75	4.74
$X_1^{\text{c}'}$	(a)	7.94	8.01	8.06	8.11	8.13	8.15
	(b)	8.33	8.26	8.22	8.18	8.17	8.16
X_3^{c}	(a)	13.27	13.28	13.29	13.30	13.31	13.31
	(b)	13.31	13.31	13.31	13.31	13.31	13.31
X_5^{c}	(a)	15.35	15.38	15.40	15.42	15.42	15.43
	(b)	15.46	15.45	15.44	15.43	15.43	15.43

[10] For example, setting this energy shift to zero changes the calculated QP energies by less than 10 meV in the case of bulk Si. Of course the QP energies always converge to the same values as in a calculation where no plane waves were added since the plane wave contribution vanishes for $N_{\text{bands}} \to \infty$.

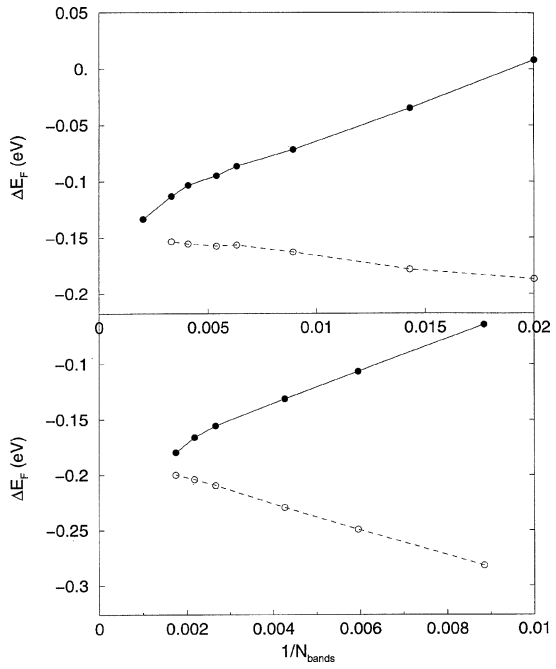

Fig. 6. Fermi energy shift $\Delta E_{\mathrm{F}} = E_{\mathrm{F}}^{\mathrm{QP}} - E_{\mathrm{F}}^{\mathrm{LDA}}$ for bulk Si (top panel) and zincblende GaN (bottom panel) as a function of the inverse of N_{bands}, the number of LDA eigenstates used for the calculation of the Green's function, calculated including (open circles, dashed lines) or not including (filled circles, solid lines) the plane-wave contribution (see text). The plane-wave contribution improves the convergence of absolute self-energies in qualitatively the same way as that of the QP energy differences shown in Figs. 4 and 5.

4.3. Tests for bulk Si and GaN

We performed a number of tests in order to assess the influence of adding the PW contribution on the convergence of self-energy and QP energies. The cutoff parameters for our tests for Si are given in Table 4. Further calculational details are as in CPC I. Fig. 4 exhibits the convergence of valence band bottom (top panel), direct gap at Γ (center), and minimal gap (bottom) as a function of the inverse of N_{bands}, the number of LDA eigenstates used for the calculation of the Green's function. Two sets of data are shown in each figure, showing results obtained with (open circles) and without (filled circles) adding the PW contribution. First of all we observe that, as expected, both sets of calculations converge to the same answer for $1/N_{\mathrm{bands}} \to 0$. However, adding the PW contribution dramatically improves convergence; the number of eigenstates needed for an accuracy of the QP energies of 30 meV is about 70% smaller when the PW contribution is added (see also Table 5). For the slowly-converging valence band bottom we find that the number of eigenstates in Eq. (2.1) can be reduced by as much as 85%.

Since bulk Si might be perceived as a particularly plane-wave like system we also tested the method for zincblende GaN. The cutoff parameters are given in Table 4, for further details of the calculation see Section 3.5 above. Fig. 5 showing the band-cutoff dependence of valence band bottom (top panel), conduction band bottom (center), and conduction state Γ_{15}^{c} (bottom) confirms that the PW contribution to the Green's function improves the band-cutoff convergence in the case of zincblende GaN, too, although not as much as for bulk Si. Adding the PW contribution decreases the number of bands needed to converge the quasiparticle energies of GaN to an accuracy of 30 meV by 30 to 50% (see also Table 6).

It can be concluded from the Fermi energy shifts $\Delta E_{\mathrm{F}} = E_{\mathrm{F}}^{\mathrm{QP}} - E_{\mathrm{F}}^{\mathrm{LDA}}$ shown in Fig. 6 that adding the PW contribution improves the convergence of absolute self-energies in qualitatively the same way as that of QP energy differences (Figs. 4 and 5).

In summary, we find that the number of LDA eigenstates (bands) needed to converge the QP energies within 30 meV can be considerably reduced by including the PW contribution described in Section 4.2 in the calculation of the Green's function.

5. Summary

In the present work we described two new features which significantly enhance the power of the real-space imaginary-time *GW* scheme for the calculation of self-energies and related quantities of solids. Fitting the smoothly decaying large-imaginary-energy/time tails and treating the remaining imaginary energy/time region numerically on a Gauss–Legendre grid allows to reduce the computational time and storage requirements of the method by a factor of seven to eight while retaining the flexibility to accomodate general functional forms of the energy dependence [11]. The tail-fitting procedure suggested in the present work turned out to be accurate and reliable. Substituting the contribution of higher unoccupied eigenstates to the Green's function Eq. (2.1) with a sum of corresponding free-electron states (plane waves) accelerates the convergence of the eigenstate sum in Eq. (2.1), thus substantially reducing the number of eigenstates and eigenvalues which have to be provided by a density-functional calculation preceding the calculation of the self-energy and simultaneously decreasing the computational effort for the calculation of the Green's function itself.

Acknowledgement

This work was supported by the Engineering and Physical Sciences Research Council, the Spain–UK Acciones Integradas program (HB 1997-011), JCyL (Grant: VA28/99) and DGES (Grant: PB95-0720). L. Reining acknowledges a grant of computer time on the C98 of IDRIS (project CP9/980544), which was used for parts of the calculations.

References

[1] L. Hedin, Phys. Rev. 139 (1965) A796.
[2] L. Hedin, S. Lundqvist, in: Solid State Physics, Vol. 23, F. Seitz, D. Turnbull, H. Ehrenreich (Eds.) (Academic, New York, 1969) pp. 1–181.
[3] M.S. Hybertsen, S.G. Louie, Phys. Rev. Lett. 55 (1985) 1418;
M.S. Hybertsen, S.G. Louie, Phys. Rev. B 34 (1986) 5390.
[4] R.W. Godby, M. Schlüter, L.J. Sham, Phys. Rev. Lett. 56 (1986) 2415;
R.W. Godby, M. Schlüter, L.J. Sham, Phys. Rev. B 37 (1988) 10159.
[5] J.E. Northrup, M.S. Hybertsen, S.G. Louie, Phys. Rev. B 39 (1989) 8198.
[6] F. Aryasetiawan, Phys. Rev. B 46 (1992) 13051.
[7] F. Aryasetiawan, O. Gunnarsson, Rep. Prog. Phys. 61 (1998) 237.
[8] H.N. Rojas, R.W. Godby, R.J. Needs, Phys. Rev. Lett. 74 (1995) 1827.
[9] M.M. Rieger, L. Steinbeck, I.D. White, H.N. Rojas, R.W. Godby, Comput. Phys. Commun. 117 (1999) 211.
[10] Z.H. Levine, S.G. Louie, Phys. Rev. B 25 (1982) 6310.
[11] S. Baroni, R. Resta, Phys. Rev. B 33 (1986) 7017.
[12] X. Blase, A. Rubio, S.G. Louie, M.L. Cohen, Bull. MRS (Fall meeting, Boston, 1994).
[13] X. Blase, A. Rubio, S.G. Louie, M.L. Cohen, Phys. Rev. B 52 (1995) R2225.
[14] N. Troullier, J.L. Martins, Phys. Rev. B 43 (1991) 1993.
[15] R. James, S.M. Woodley, Solid State Commun. 97 (1996) 935.
[16] A. Fleszar, W. Hanke, Phys. Rev. B 56 (1997) 10228.

[11] The CPU time (on a Digital Alpha 500/500 workstation), disk space and memory required for computing the full self-energy and 48 self-energy matrix elements for Si (GaN) are 35 minutes (194 minutes), 243 MB (1213 MB) and 65 MB (123 MB), respectively, with the parameters given in Table 1.